

AMENDMENTS TO THE SPECIFICATION:

Please replace the paragraph beginning on page 3, line 24 with the following amended paragraph:

In the above at first cited article, prosodic modeling is mainly based on only very local features, whereby for each inter-word boundary prosodic features of the word immediately preceding and following the boundary, or alternatively within an empirically optimized window of 20 frames before and after the boundary, are analyzed. In particular, prosodic features are extracted that reflect pause durations, phone durations, [[pitch]] pitch information, and voice quality information. Pause features are extracted at the inter-word boundaries. Pause duration, a fundamental frequency (F0), and voice quality features are extracted mainly from the word and window preceding the boundary. In addition, pitch-related features reflecting the difference in pitch across the boundary are included in the analysis.

Please replace the paragraph beginning on page 17, line 24 with the following amended paragraph:

A further prosodic feature is the difference F2 – F1. It has been found that a speaker after a segment boundary in most cases does not continue with the same pitch, thus resulting in a so-called pitch so-called pitch reset.

Please replace the paragraph beginning on page 20, line 2 with the following amended paragraph:

In Fig. 5A, a digitized speech signal 600 is input to an F0 processor depicted in Fig. 3 that computes 610 a continuous F0 data from the speech signal. Only by the criterion F0 = ON/OFF, as described beforehand, the speech signal is presegmented 620 into speech segments. For each segment 630 it is evaluated 640 whether F0 is defined or [[no]] not defined. In the case of a not defined F0 (i.e. F0 = OFF) a candidate segment boundary is

assumed as described above and, starting from that boundary, prosodic features will be computed 650. The feature values are input into a classification tree ([[s.]] see Fig. 6) and each candidate segment is classified thereby revealing, as a result, the existence or non-existence of a semantic or syntactic speech unit.

Please replace the paragraph beginning on page 21, line 7 with the following amended paragraph:

The Compute features step 650 shown in Fig. 5A is now depicted in more detail referring to Fig. 5C. In a first step 900, starting from a candidate boundary with F0 = OFF, F0 itself is used as a prosodic features feature and computed accordingly. [[IN]] In a further step 910, prosodic features in a time window lying before the candidate boundary are computed. In a next step 920, prosodic features are computed in a time window lying after the candidate boundary.

Please replace the paragraph beginning on page 22, line 18 with the following amended paragraph:

It is noted again hereby that, for the approach according to the present invention, there is no need for a speech recognizer to be invoked in the segmentation process.